

## VU Research Portal

### **Effecten van gedifferentieerd wiskundeonderwijs: de toepassing van een model voor multilevelanalyse bij curriculumevaluatie.**

Terwel, J.; Van den Eeden, P.

#### ***published in***

Tijdschrift voor Onderwijsresearch  
1990

#### ***document version***

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

#### ***citation for published version (APA)***

Terwel, J., & Van den Eeden, P. (1990). Effecten van gedifferentieerd wiskundeonderwijs: de toepassing van een model voor multilevelanalyse bij curriculumevaluatie. *Tijdschrift voor Onderwijsresearch*, 15(5), 273-284.

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

#### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

# Effecten van gedifferentieerd wiskundeonderwijs: de toepassing van een model voor multilevelanalyse bij curriculumevaluatie

J. Terwel en P. van den Eeden

*Rijksuniversiteit Utrecht, Vakgroep Onderwijskunde\* en Vrije Universiteit, Vakgroep Methoden en Technieken van Sociaal-Wetenschappelijk Onderzoek\*\**

## ABSTRACT

The application of a model of multilevel analysis in the case of curriculum evaluation: effects of differentiated instruction in mathematics.

The aim of this article is to describe the effects of a newly developed model of instruction in mathematics in the second grade of Dutch comprehensive schools (LBO-MAVO-HAVO-VWO). Main characteristics of the model are: 'cooperative learning' and 'mathematics in real life situations' (contexts). The database was adopted from an earlier study of Terwel, Herfs, Dekker and Akkermans (1988), and comprises data of 482 pupils in 22 classes. Application of the random coefficient model of multilevel analysis shows that there are significant effects of class-composition. If class-composition is omitted from the model, curriculum-implementation shows a significant effect. No effect of off-task behavior was found. The article shows the fruitfulness of the use of the random coefficient model of multilevel analyses for the assessment of effects in evaluation research.

## VRAAGSTELLING

Bij de aandacht voor de totstandkoming van leerresultaten zijn er tenminste drie richtingen te onderscheiden. In de eerste richting ligt het accent op de samenhang tussen leerresultaten en de recrutering van leerlingen. In de tweede richting tracht men leerresultaten te verklaren door de kwaliteit van het curriculum en de instructie.

Daarnaast bestaat er nog een derde richting van aandacht. Die richt zich op de invloed die het onderwijssysteem zelf uitoefent op het leerproces van de leerlingen. Een algemeen centraal element in het onderwijssysteem is dat het is georganiseerd in klassen. Het leerproces zelf vindt bij de leerlingen plaats. Met andere woorden: het onderwijssysteem bezit typisch een geneste structuur van meerdere levels. We spreken van "levels" om een geneste structuur aan te duiden, dit ter onderscheiding van de term "niveaus" in andere betekenissen. In dit artikel, dat zich richt op de situatie in het voortgezet onderwijs, worden twee levels in de analyse betrokken: leerlingen en klassen. In dit verband is de vraag van belang of het leerproces binnen elk van deze levels anders verloopt, en of het antwoord op die vraag verschilt per niveaugroep (zwakke, middelmatige en sterke leerlingen). Het grondidee hierbij is dat de niveaugroep waartoe een leerling behoort en zijn klas een specifieke inhoud kunnen geven aan de leeromgeving van een leerling. Daarmee is gezegd dat de leerresultaten van een leerling mede afhankelijk kunnen zijn van niveaugroep en klas. Het doel van dit artikel is nagaan in hoeverre deze vraag met behulp van empirisch materiaal kan worden beantwoord.

\* Vakgroep Onderwijskunde, Postbus 80140, 3508 TC Utrecht.

\*\* Vakgroep Methoden en Technieken van Sociaal-Wetenschappelijk Onderzoek, Koningslaan 22-24, 1075 AD Amsterdam.

De algemene hypothese voor dit artikel is dat het leren van individuele leerlingen onder invloed staat van variabelen op hogere levels, in dit geval de klas. We verwachten dat het individueel leerproces niet een op zichzelf staand gebeuren is, maar afhankelijk is van kenmerken van de klas. Leerlingen van verschillend niveau zullen naar verwachting niet in gelijke mate profiteren van de voordelen van een gunstige classesamenstelling, de curriculuminnovatie en de taakgerichtheid van de klas.

De analyses worden voor elke niveaugroep afzonderlijk uitgevoerd. Men zou kunnen denken dat het hier gaat om drie levels, met de niveaugroep als hoogste level. Dat is evenwel een misvatting, omdat met de drie niveaugroepen niet evenzoveel eenheden zijn die een steekproef uit een of andere populatie vormen, maar wel categorieën van een variabele zijn (en wel behorend bij de leerling). Aan het te hanteren random coëfficiënt model ligt noodzakelijkerwijze een design van getrapte steekproeftrekking ten grondslag, dat niet zou worden gerealiseerd indien de niveaugroep als hoogste level zou worden beschouwd. In een dergelijke situatie wordt wel aangenomen dat er kan worden gegeneraliseerd naar een voorshands onbekende theoretische populatie. Dit gebeurt bijvoorbeeld in onderzoek naar cohorteffecten op schoolverschillen van loopbanen van leerlingen. Een dergelijke aanname van een theoretische populatie van alle mogelijke niveaugroepen is wellicht in principe mogelijk, maar lijkt in dit geval wat ver gezocht. Door een gescheiden analyse van de verschillende niveaugroepen is het in elk geval mogelijk de niveaugroepen apart te karakteriseren.

## DATA

De analyse is uitgevoerd met behulp van een dataset, afkomstig van het project Interne Differentiatie (ID 12-16) (Terwel, Herfs, Dekker en Akkermans, 1988). Dat project was gericht op het vaststellen van de effectiviteit van een curriculum voor het wiskundeonderwijs dat is gebaseerd op wiskundig-didactische principes van Freudenthal (1973 a en b, en 1980). Deze principes zijn:

- a. klassikale introducties en reflecties in heterogene klassen;
- b. samenwerking in heterogene subgroepjes;
- c. niveaus in het leerproces (verschillen in oplossingen);
- d. wiskunde in rijke contexten.

Het geheel van deze principes duiden we aan met 'het (wiskunde didactisch) model Freudenthal'. Het curriculum is ontwikkeld door het Instituut voor de Leerplanontwikkeling (SLO).

De gegevens werden verzameld in drie brede scholengemeenschappen, van 482 leerlingen en 12 leraren in 22 tweede klassen, de adviezen van de leerlingen varieerden van LBO tot en met VWO. De voormeting vond plaats in oktober 1985 en de nameting in mei 1986. Naast de gegevens die betrekking hebben op de leerling, zijn ook gegevens over de klas en de niveaugroep in de analyse betrokken.

## HYPOTHESEN

Achtergrond voor de hieronder geformuleerde specifieke hypothesen is de verwachting dat de leerwinst onder invloed staat van ten minste twee levels: de leerling en de klas, en dat deze invloed per niveaugroep zal verschillen.

Onze hypothesen zijn te verdelen in twee groepen.

1. De eerste hypothese is dat er een zelfstandig effect is van de voortoets (TOT1) op de natoets (TOT3). Het resultaat van de voormeting is rechtstreeks van invloed op de nameting. Dit individueel effect, op het level van de leerling, duiden we aan met de term "leerwinst". Vandaar dat we die hypothese aanduiden als de leerwinsthypothese. Het kan zijn dat deze leerwinst een



intra-individueel proces is, maar het is daarnaast mogelijk dat zij afhankelijk is van een aantal kenmerken in de onderwijssituatie. Te denken valt hier aan kenmerken gelegen in de klas en kenmerken van de niveaugroep waartoe de leerling behoort.

2. De tweede categorie hypothesen is gelegen op het level van de klas. Deze klassehypothesen nuanceren de leerwinsthypothese door de kenmerken van de klassesituatie erbij te betrekken. Het betreft drie hypothesen: de samenstellingshypothese, de implementatiehypothese, en de hypothese over taakgerichtheid van de klas.

2a. De samenstellingshypothese luidt: hoe hoger het gemiddelde beginniveau van de klas des te hoger is de score op de eindtoets. De variabele classesamenstelling is bepaald door het gemiddelde van de klas op de voortoets (MTOT). De variabele MTOT betreft het level van de klas aan het begin van het schooljaar. Voor de formulering van deze hypothese is aangesloten bij onderzoek naar het effect van klassecompositie op het leren van individuele leerlingen (vgl. Beckermann en Good 1981, Good en Marshall 1984, Oakes 1986, Dar en Resh 1986). Een grondgedachte hierbij is dat het percentage sterke leerlingen in een klas een positief effect heeft op het leren van alle leerlingen in die klas. In klassen met veel zwakke leerlingen verloopt het onderwijsleerproces minder goed en er wordt minder leerstof behandeld dan in klassen met veel sterke leerlingen (vgl. Dreeben en Barr 1987, 34).

2b. De implementatiehypothese luidt: hoe meer het curriculum volgens de bedoelingen is uitgevoerd des te hoger zijn de scores op de eindtoets. De variabele MPERCIA betreft de mate waarin het curriculum in de klas volgens de bedoelingen van de leerplanontwikkelaars is uitgevoerd. MPERCIA is vastgesteld aan de hand van het gemiddelde per klas van de perceptiescore van het curriculum door de leerlingen. De leerlingpercepties zijn gemeten met een vragenlijst voor de perceptie van de onderwijsleersituatie, de zogenoemde PERCIA-schaal (PERceptie van het Curriculum In Actie). Deze vragenlijst voor leerlingen betreft een curriculum-specifieke versie van schalen voor het bepalen van de perceptie van de leeromgeving. Leerlingpercepties zijn als zodanig op te vatten als indicatoren van implementatie. Bovendien zijn er indicaties uit onderzoek dat de gemiddelde perceptie van leerlingen vrij hoog correleert met geobserveerd gedrag van leraren (Fraser en Tobin 1989). Overigens gaat onze voorkeur uit naar observatiegegevens betreffende de implementatie; deze zijn echter niet beschikbaar. Het is niet uit te sluiten dat in de leerlingpercepties houdingsaspecten meespelen. Datgene wat een leerling waarneemt kan beïnvloed zijn door zijn houding ten opzichte van het wiskunde-onderwijs of door zijn specifieke leerbehoeften.

In de PERCIA-vragenlijst zijn verschillende kenmerken van het nieuwe wiskundeonderwijs opgenomen. Deze kenmerken betreffen de leiding en de begeleiding door de leraar, de samenwerking in groepjes, de taakgerichtheid en het didactische principe 'wiskunde in contexten'. Bij keuze en uitwerking van deze begrippen is aansluiting gezocht bij de vakdidactiek, bij leertheorieën en bij resultaten van empirisch onderzoek betreffende effectiviteit van instructie (vgl. Freudenthal 1973 en 1980, Terwel 1984 en 1989).

2c. De taakgerichtheidshypothese is: hoe meer niet-taakgericht gedrag in de klas des te lager de scores op de eindtoets (en omgekeerd) (vgl. Brophy, 1986). We verwachten in alle gevallen een negatief effect op de eindtoetsscores. De variabele NTAJOB is het percentage niet-taakgericht gedrag van de leerlingen in de klas. Deze variabele is gemeten aan de hand van systematische, kwantitatieve observaties in de klas.

Het behoren tot een niveaugroep heeft mogelijk ook invloed op de leerwinst van de leerlingen. We verwachten dat de leerwinst hoger is naarmate een leerling tot een hogere niveaugroep behoort. Het onderwijs volgens het 'model Freudenthal' veronderstelt wellicht kennis en vaardigheden die sterke leerlingen reeds bezitten waardoor zij meer van het onderwijs kunnen profiteren. Bovendien veronderstellen wij dat de sterke leerlingen relatief weinig beïnvloed worden door negatieve factoren in de leeromgeving. Er zijn indicaties uit onderzoek en theorievorming van Dar en Resh (1986) dat leerlingen uit de middelste niveaugroep gevoeliger zijn voor de factor klasse-samenstelling dan sterke leerlingen. Op basis van de verdeling van TOT1 is er een indeling gemaakt in drie niveaugroepen, laag, midden en hoog, conform eerder onder-

zoek (Terwel, Herfs, Dekker en Akkermans 1988). Leerlingen uit een bepaalde niveaugroep ontvangen onderwijs in verschillende klassen. Tussen deze klassen bestaan verschillen in de samenstelling dat wil zeggen verschillen in gemiddelde voorkennis (gemiddelde TOT1). We gaan na of er verschillen tussen de niveaugroepen bestaan in gevoeligheid voor verschillen in MTOT, MPERCIA en NTAKOB.

Voor een beschrijving van de onderzoeksinstrumenten en hun betrouwbaarheid en validiteit verwijzen we naar het oorspronkelijke onderzoeksverslag (Terwel, Herfs, Dekker en Akkermans 1988).

### HET MODEL

Het uitgangspunt van onze analyse is de regressie van de natoets (TOT3) op de voortoets (TOT1). Over de effecten van klassevariabelen op deze regressie van TOT3 op TOT1 is hierboven een aantal hypothesen geformuleerd. Deze hypothesen worden getoetst binnen het random coëfficiënt (RC) model voor multilevelanalyse. Dit model is opgebouwd uit twee stappen: in de eerste stap is er sprake van een binnengroepregressie. In de tweede stap worden de uitkomsten daarvan ingevoerd in een tussengroepregressie (De Leeuw en Kreft, 1986, Aitkin en Longford, 1986, Raudenbush en Bryk, 1986, en Goldstein, 1987). In de meest eenvoudige vorm bevat het model twee levels en kan het model in twee stappen worden uitgeschreven, hoewel de toetsing simultaan plaatsvindt.

De twee stappen van het model kunnen als volgt worden uitgeschreven, waarbij hier als groepslevel de klas wordt genomen. In de eerste stap wordt de binnengroepregressie uitgedrukt in de vergelijking:

$$\text{TOT3}_{ij} = a_j + b_j \text{TOT1}_{ij} + z_{ij}$$

waarin:

- i: de individuele leerling ( $i = 1 \dots I$ )
- j: de klas ( $j = 1 \dots J$ )
- $a_j$ : het intercept voor klas j
- $b_j$ : de helling voor klas j
- $z_{ij}$ : de storingsterm met variantie  $s^2$ .
- $\text{TOT3}_{ij}$ : de wiskundescore einde studiejaar.
- $\text{TOT1}_{ij}$ : wiskundescore begin studiejaar.

In de tweede stap is er sprake van een tussengroepregressie van het intercept  $a_j$  en de helling  $b_j$  op de klassevariabelen.

$a_j = A_0 + A_1 \text{MTOT}_j + A_2 \text{MPERCIA}_j + A_3 \text{NTAKOB}_j + d_j$   
 $b_j = B_0 + B_1 \text{MTOT}_j + B_2 \text{MPERCIA}_j + B_3 \text{NTAKOB}_j + e_j$   
 met  $d_j$  en  $e_j$  de storingstermen behorend bij respectievelijk  $a_j$  en  $b_j$  met varianties  $t^2$  en  $u^2$ .  $A_0$  staat voor een algemene constante,  $A_m$  voor het direct effect van een klasvariabele M,  $B_0$  voor het direct individueel effect en  $B_m$  voor het effect van een klasvariabele op het effect van TOT1<sub>ij</sub> op TOT3<sub>ij</sub>.

Het RC-model (De Leeuw en Kreft, 1986) staat overigens ook bekend als "variance decomposition-model" (Aitkin en Longford, 1986) en "hierarchical linear model" (Raudenbush en Bryk, 1986). Het vertoont overeenkomst met de "separate equation approach" van Boyd en Iversen (1979) en het model van de "tweestaps-procedure" (Van den Eeden en Saris, 1984), in zoverre de coëfficiënten worden geschat. Met behulp van dit model is het mogelijk de klasspecifieke onderwijskansen af te lezen aan de verschillen in de intercepten en de hellingen van een regressie waarin de invloed van TOT1 van leerlingen op hun TOT3 tot uiting wordt gebracht. Bij de analyses is gebruik gemaakt van het VARCL-programma van Longford (1986).

## ANALYSE

Alvorens in te gaan op de multilevelanalyse, vermelden we in de Tabellen 1 en 2 de verdelingskenmerken van respectievelijk de individuele en de klassevariabelen.

Bij de toetsing zijn meerdere modellen gehanteerd. In Model 1 wordt alleen de variantie van de afhankelijke variabele (TOT3) vermeld. In Model 2 wordt aangegeven welk deel van deze variantie aan de leerling en welk deel aan de klas kan worden toegeschreven. Met behulp van Model 3 wordt vastgesteld in hoeverre de leerlingvariantie door de voortoets (TOT1) wordt gebonden, in overeenstemming met de leerwinsthypothese. In Model 4 wordt vervolgens nagegaan in hoeverre door de introductie van de variabele TOT1 er een tussenklasvariantie rond de regressie van TOT3 op TOT1 aanwezig is. In Model 5 wordt gekeken in hoeverre de eventuele, laatstgenoemde tussenklasvariantie wordt verklaard door de klassevariabelen MTOT, MPER-

Tabel 1. Kenmerken van de verdelingen van TOT1 en TOT3 en de regressiecoëfficiënten van TOT3 op TOT1 per niveaugroep. Tussen haken is de standaardfout vermeld.

		TOT3	TOT1	toename	regressiecoëfficiënten
NIVEAU-GROEP 1	gemiddelde	29.54	20.62	8.82	a = 8.74 ( 2.34)
	st.dev	8.98	4.88		b = 1.01 ( .11)
	min	2	4.5		R <sup>2</sup> = .30
	max	54	26.5		
	cases	197			
	klassen	22			
NIVEAU-GROEP 2	gemiddelde	41.62	30.20	11.42	a = 1.39 (11.19)
	st.dev	10.87	2.26		b = -.34 ( .37)
	min	9	27		R <sup>2</sup> = .08
	max	75	34.5		
	cases	158			
	klassen	21			
NIVEAU-GROEP 3	gemiddelde	54.93	40.91	14.03	a = 9.68 ( 5.88)
	st.dev	11.40	5.88		b = 1.11 ( .14)
	min	32	35		R <sup>2</sup> = .33
	max	86	63		
	cases	127			
	klassen	18			

Tabel 2. Verdelingskenmerken van de klasse-variabelen, alsmede de correlaties tussen de klasse-variabelen.

	gemiddelde	standaard deviatie	minimum	maximum
MTOT1	27.804	5.721	18.231	39.032
MPERCIA	3.305	.206	2.881	3.605
NTAKOB	25.597	24.027	.370	74.620
MTOT1	1.000			
MPERCIA	.570	1.000		
NTAKOB	-.676	-.525	1.000	
	MTOT1	MPERCIA	NTAKOB	



CIA en NTAKOB, overeenkomstig onze gerichtheid op de drie klassehypotheseën. In Model 6, tenslotte, wordt vastgesteld in hoeverre de eventueel resterende interceptvarianties eveneens door deze klassevariabelen worden gebonden. Bij deze toetsingsprocedure wordt telkens gekeken of de modelpassings-parameters DEV van opeenvolgende modellen elkaar verschillen, daarbij rekening houdend met het verschil in aantal vrijheidsgraden. Niet-significante coëfficiënten worden op nul gefixeerd. In alle gevallen is het 95 procents-niveau aangehouden.

We gaan nu in op de gegevens uit de multilevelanalyse. De uitkomsten vermelden de ongestandaardiseerde coëfficiënten. Voor de interpretatie daarvan dient men rekening te houden met de verdelingskenmerken als gemiddelde, minimum en maximum (zie Tabel 1 en Tabel 2). In Tabel 3 zijn de resultaten van de multilevelanalyses vermeld. Niet-relevante uitkomsten zijn weggelaten.

In Model 1 zijn alle vaste coëfficiënten en de random coëfficiënten op groepsniveau weggelaten. Het resultaat is dat alleen de totale variantie van de afhankelijke variabele,  $s^2$ , wordt vermeld. Het valt op in Model 1, dat de te verklaren variantie verschilt per niveaugroep. Zij wordt groter naarmate de niveaugroep hoger is.

In Model 2 wordt de totale variantie van TOT3 uit Model 1 gesplitst in een individueel deel en een klasse-deel (weergegeven door  $s^2$  en  $t^2$ ), of anders gezegd, in een binnengroepsvariantie en een tussengroepsvariantie. Bij de twee laagste niveaugroepen is er een splitsing van een

Tabel 3. Uitkomst van de multilevelanalyse. Niet-relevante uitkomsten zijn weggelaten. Tussen haakjes is de standaardfout vermeld.

NIVEAUGROEP 1 (Zwakke leerlingen)						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
fixed deel						
individueel effect						
$B_0$ effect TOT1			.91 (.11)	.91 (.11)	.39 (.11)	.85 (.11)
klasse effect						
interceptvariantie						
verklaard door						
$A_1$ MTOT						.47 (.15)
$A_2$ MPERCIA						
$A_3$ NTAKOB						
hellingvariantie						
verklaard door						
$B_1$ MTOT					.02 (.01)	
$B_2$ MPERCIA						
$B_3$ NTAKOB						
random deel						
$s^2$ individueel	80.28	63.05*	48.24	48.24	47.97	47.71
$t^2$ intercept (klas)	—	17.89*	8.06	8.06	4.81	4.50
$v^2$ helling (klas)						
$R^2$ individueel			23.5	23.5	23.9	24.3
$R^2$ intercept			44.7	44.7	73.1	74.9
$R^2$ helling						
DEV	1423.000	1401.623	1341.757	1341.757	1335.056	1333.400
DEV-verschil		21.377	59.866	0	6.701	8.357
DF-verschil		1	1	1	1	1
met Model		1	2	3	3	4

\* Dit absolute getal is uitgangspunt voor berekening van de verklaarde variantie in de volgende modellen.

(Vervolg Tabel 3)

NIVEAUGROEP 2 (Middelmatige leerlingen)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
fixed deel						
individueel effect						
B <sub>0</sub> effect TOT1			1.03 (.34)	1.03 (.34)		.90 (.19)
klasse effect						
interceptvariantie verklaard door						
A <sub>1</sub> MTOT						.98 (.19)
A <sub>2</sub> MPERCIA						
A <sub>3</sub> NTAKOB						
hellingvariantie verklaard door						
B <sub>1</sub> MTOT					.31 (.01)	
B <sub>2</sub> MPERCIA						
B <sub>3</sub> NTAKOB						
random deel						
s <sup>2</sup> individueel	117.303	87.177*	84.104	81.104	84.416	84.360
t <sup>2</sup> intercept (klas)		38.609*	30.230	30.230	5.877	5.086
v <sup>2</sup> helling (klas)						
R <sup>2</sup> individueel			3.7	3.7	3.2	3.3
R <sup>2</sup> intercept			21.8	21.8	84.8	86.8
R <sup>2</sup> helling						
DEV	1201.216	1183.490	1174.712	1174.712	1157.784	1156.742
DEV-verschil		17.726	8.778	0	16.928	17.970
DF-verschil		1	1	0	1	1
met Model		1	2	3	3	4

NIVEAUGROEP 3 (Sterke leerlingen)

fixed deel						
individueel effect						
B <sub>0</sub> effect TOT1			1.11 (.14)	1.11 (.14)	1.11 (.14)	1.11 (.14)
klasse effect						
interceptvariantie verklaard door						
A <sub>1</sub> MTOT						
A <sub>2</sub> MPERCIA						
A <sub>3</sub> NTAKOB						
hellingvariantie verklaard door						
B <sub>1</sub> MTOT						
B <sub>2</sub> MPERCIA						
B <sub>3</sub> NTAKOB						
random deel						
s <sup>2</sup> individueel	128.844	128.844*	86.874	86.874	86.874	86.874
t <sup>2</sup> intercept (klas)						
v <sup>2</sup> helling (klas)						
R <sup>2</sup> individueel			32.6	32.6	32.6	32.6
R <sup>2</sup> intercept						
R <sup>2</sup> helling						
DEV	977.453	977.453	927.397	927.397	927.397	927.397
DEV-verschil		0	50.056	0	0	0
DF-verschil		1	1	0	0	0
met Model		1	2	3	3	4



individuele variantie en een klassevariantie van de afhankelijke variabele TOT3. De klassevariantie is voor niveaugroep 1 en 2 respectievelijk 22.1 en 30 percent. Bij de hoogste niveaugroep is er geen significante tussenklassen-variantie. De prestaties van de leerlingen uit deze groep lijken onafhankelijk te zijn van klasse-invloeden.

In Model 3 wordt behalve de random coëfficiënten  $s^2$  en  $t^2$  de vaste coëfficiënt  $B_0$  geschat;  $s^2$  representeert dan de restvariantie op het niveau van de leerling. Model 3 stelt ons in staat de leerwinsthypothese over het effect van de voortoets (TOT1) te toetsen. De leerwinsthypothese voor de individuele leerlingen geldt in elke niveaugroep, zij het dat er opmerkelijke verschillen optreden. In de laagste niveaugroep verklaart de variabele TOT1 23.5 percent van de individuele variantie en bijna de helft van de tussenklassenvariantie (44.7 percent). In de middelste niveaugroep verklaart TOT1 slechts 3.7 percent van de individuele variantie en 21.8 percent van de tussenklasvariantie. In de hoogste niveaugroep is het percentage verklaarde individuele variantie aanzienlijk hoger dan in de twee lagere niveaugroepen: 32.6 percent. De score op de natoets van de leerlingen uit de middelste niveaugroep lijkt nog het minst afhankelijk te zijn van hun score op de voortoets.

In Model 4 wordt vervolgens de random coëfficiënt voor de tussenklasvariantie van de helling van de regressie van TOT3 op TOT1,  $v^2$ , aan het model toegevoegd. In alle niveaugroepen is gebleken dat er in Model 4 geen variantie is tussen de hellingen.

Deze uitkomst impliceert niet dat Model 5 niet voor toetsing in aanmerking komt. In Model 5 wordt vastgesteld welk deel van de tussenklasvariantie van de hellingen wordt verklaard door de klassevariabelen MTOT, MPERCIA en NTAKOB. Dit betekent dat de coëfficiënten  $B_1$ ,  $B_2$  en  $B_3$  worden geschat. Bij het onderzoek naar een afwijking van Model 5 ten opzichte van Model 3 zijn we evenwel in een betere uitgangspositie dan bij een dergelijk onderzoek van Model 4. We gebruiken hier theoretische informatie die leert dat de eventuele variatie in hellingen niet toevallig is, maar samenhangt met bepaalde klassevariabelen. De power van een toets van Model 3 (de nulhypothese) tegen Model 5 (de alternatieve hypothese) is dus in het algemeen hoger dan de power van een toets van Model 3 tegen Model 4.

Zoals gezegd, is er evenwel bij geen der niveaugroepen een significante variantie aanwezig van de regressiehellingen van TOT3 op TOT1 tussen de klassen. De verschillen zijn vooral aanwezig tussen de niveaugroepen, maar niet tussen de klassen binnen elke niveaugroep. De toetsing van Model 5 tegen Model 3 evenwel laat in de laagste en de middelste niveaugroep zien dat de variantie in de hellingen tussen de klassen samenhangt met MTOT; in de hoogste niveaugroep is deze samenhang afwezig. Het verband is het sterkst bij de middelste niveaugroep. Dit betekent dat de bovengenoemde hypothesen over de afhankelijkheid van de leerwinst van de klassevariabelen alleen met betrekking tot de classesamenstelling kan worden bevestigd.

In Model 6 toetsen we de drie klassehypothesen. Dit gebeurt door naast de geaccepteerde coëfficiënten ook de coëfficiënten  $A_1$ ,  $A_2$  en  $A_3$  in het model ter schatting op te nemen. Alleen bij de twee laagste niveaugroepen leerlingen hangt het gemiddelde van de TOT3-scores van een klas samen met de gemiddelde TOT1-score (MTOT). Dit verband is niet te begrijpen als een gevolg van een aggregatie van het zojuist genoemde verband tussen TOT1 en TOT3 op het individuele level, omdat daarvoor reeds is gecontroleerd bij het toetsen van de leerwinsthypothese. Toen bleek (in Model 3) dat van de variantie in de gemiddelde TOT3 in de eerste niveaugroep 45 percent werd verklaard door introductie van TOT1 en in de middelste niveaugroep voor 21.8 percent. Bij de leerlingen uit de hoogste niveau-groep is de cross-level interactie tussen TOT1 en MTOT in hun effect op TOT3 evenwel afwezig. De andere klassevariabelen MPERCIA en NTAKOB hebben geen significant effect op TOT3.

Dit resultaat is niet strijdig met de uitkomst van Model 5. Zowel Model 5 als Model 6 blijkt in voldoende mate bij de oorspronkelijke gegevens te passen. Omdat wij onze hypothesen hebben ontwikkeld met betrekking tot Model 5 en omdat de percentages verklaarde varianties niet belangrijk van elkaar verschillen geven we aan de uitkomst daarvan de voorkeur.

De bovenstaande uitkomsten laten zien welke verklaring er door het gehanteerde RC-model wordt bereikt. De verhouding tussen de individuele en tussenklassen-variantie is in de twee

laagste niveaugroepen ongeveer 3 : 1 respectievelijk 77.9 versus 22.1 en 74.3 versus 25.7. Bij de laagste niveaugroep is de verklaring van deze varianties in een verhouding van 1 : 3 (dat wil zeggen 24.3 versus 74.9) van de individuele en de tussenklassen-variantie en in de middelste niveaugroep ongeveer 1 : 30 (dat wil zeggen 3.3 versus 86.8). De verschillen in deze niveaugroepen manifesteren zich in de verschillen in de verklaring op de twee betrokken levels. De middelste niveaugroep is in sterke mate afhankelijk van de klasseomgeving. In de middelste niveaugroep is de coëfficiënt voor MTOT .98, in de laagste niveaugroep is deze coëfficiënt .47 terwijl deze coëfficiënt in de hoogste niveaugroep niet significant verschilt van nul. In een contrast hiermee staat de hoogste niveaugroep, waar alleen op het individueel level een verklaring plaatsvindt, die echter sterker is dan in de andere niveaugroepen.

De implementatiehypothese bleek in eerste instantie, in het door ons gekozen model, niet bevestigd te worden. Er bleek echter een samenhang te bestaan tussen MTOT en MPERCIA (zie Tabel 2). Dit duidt op collineariteit waardoor de implementatiehypothese in een model waarin ook MTOT voorkomt niet adequaat kan worden getoetst. Daarom hebben we ook een analyse met Model 6 uitgevoerd naar het effect van MPERCIA zonder dat daarin MTOT was opgenomen. In deze laatste analyse was er wel een significant effect van MPERCIA op de helling, hetgeen overeenkomt met de verwachtingen en de bevindingen van de eerdere studie. In Tabel 4 zijn de uitkomsten vermeld.

Tabel 4. Uitkomst van de multilevelanalyse met weglating van de variabele MTOT in Model 6, voor alle niveaugroepen. De vergelijking vindt plaats met Model 4.

	Niveaugroep 1	Niveaugroep 2	Niveaugroep 3
fixed deel			
individueel effect			
B <sub>0</sub> effect TOT1	.87 (.11)	1.01 (.34)	
klasse effect			
interceptvariantie			
verklaard door			
A <sub>2</sub> MPERCIA			
A <sub>3</sub> NTAKOB			
hellingvariantie			
verklaard door			
B <sub>2</sub> MPERCIA	8.22 (3.55)	15.56 (6.56)	
B <sub>3</sub> NTAKOB			
random deel			
s <sup>2</sup> individueel	47.857	83.814	86.874
t <sup>2</sup> intercept (klas)	5.923	21.470	
v <sup>2</sup> helling (klas)			
R <sup>2</sup> individueel	24.1	4.9	32.6
R <sup>2</sup> intercept	66.1	44.4	
R <sup>2</sup> helling			
DEV	1336.734	1169.514	927.397
DEV-verschil	5.023	5.198	0
DF-verschil met Model 4	1	1	1

De uitkomsten in Tabel 4 duiden op een effect van de implementatie van het Model Freudenthal in de laagste en middelste niveaugroepen. Er is geen effect in de hoogste niveaugroep. De sterke leerlingen lijken weinig gevoelig te zijn voor verschillen in implementatie, d.w.z. voor de kwaliteit van de leeromgeving in de klas.

## CONCLUSIES

In dit artikel zijn vier hypothesen getoetst tegen de achtergrond van de verwachting dat het leerproces van leerlingen onder invloed staat van variabelen op het hogere level in casu de klas. In het verlengde van deze vier hypothesen zijn de volgende hoofdconclusies te trekken.

De leerwinsthypothese leverde op dat er een zelfstandig effect blijkt te bestaan van de voortoets op de natoets. Datgene wat leerlingen aan wiskundige voorkennis bezitten blijkt rechtstreeks van invloed op hun prestaties na het doorlopen van het curriculum. Deze hypothese is geldig voor elke niveau-groep. De volgende conclusies vullen de conclusie over de leerwinst-hypothese aan door het effect van klassevariabelen erbij te betrekken.

Voor de samenstellingshypothese geldt het volgende: De samenstelling van de klas, tot uitdrukking gebracht door het klassegemiddelde, blijkt het verwachte effect te hebben op de leerwinst van de zwakke en middelmatige leerlingen. Hoe hoger het klassegemiddelde hoe beter hun score op de voortoets wordt omgezet in hun score op de natoets (zie hellingseffect, Tabel 3, Model 5). Daarnaast geldt dat bij de zwakke en middelmatige leerlingen de samenstelling van de klas effect heeft op hun prestaties. Hoe hoger het klassegemiddelde hoe hoger de prestaties van de leerlingen (zie effect op het intercept, Tabel 3, Model 6). Omdat het hier gaat om twee verschillende interpretaties en de eerste ons het meest plausibel voorkomt, geven we de voorkeur aan de eerste. In beide gevallen geldt dat de middelmatige leerlingen het meest gevoelig blijken voor de samenstelling van de klas. Voor de sterke leerlingen gaan de hypothesen echter niet op. De leerlingen in de hoogste niveaugroep lijken individualistisch georiënteerd. Hun prestaties worden niet beïnvloed door de samenstelling van de klas.

De implementatiehypothese bleek aanvankelijk, in het door ons gekozen model, niet bevestigd te worden. Na weglating van de variabele classesamenstelling uit het model (ter vermijding van collineariteit) bleek er echter toch een significant effect van de implementatie van het Model Freudenthal, waarmee de implementatiehypothese en de bevindingen uit de eerdere studie zijn bevestigd.

De hypothese over de taakgerichtheid van de klas blijkt niet bevestigd te worden, ook niet na weglating van MTOT. In de door ons gehanteerde modellen is geen effect aangetoond van de taakgerichtheid van de klas.

## DISCUSSIE

We besteden eerst aandacht aan de conclusies betreffende het effect van classesamenstelling. De prestaties van de leerlingen in de hoogste niveaugroep zijn relatief onafhankelijk van de samenstelling van de klas. De twee andere niveaugroepen laten in gelijke mate individuele en klasseverschillen zien. In de middelste niveaugroep kunnen die individueel nauwelijks worden verklaard, maar door de classesamenstelling des te meer. De classesamenstellingshypothese is dus vooral van belang in de middelste niveaugroep.

Interpretaties van deze bevindingen kunnen worden ontleend aan Dar en Resh (1986) en aan Fend (1982). Dar en Resh verklaren de grotere gevoeligheid van de middengroep op basis van een drempelhypothese. Om te kunnen profiteren van een verrijkte leeromgeving moet men een minimum aan kennis en vaardigheden bezitten. Het is juist de middengroep die over de startcompetenties beschikt om de sprong naar een klas met een hoger niveau te kunnen maken. Zij winnen relatief veel in een hogere klas en verliezen overeenkomstig in een zwakke klas. Fend wijst op een tweede interpretatiemogelijkheid, die hij evenwel op schoolniveau heeft geformuleerd. Hij noemt het "differentiellen Förderungseffekt" als een van de belangrijkste en meest consistente gegevens uit het vergelijkende onderzoek naar de Gesamtschule. Dit effect betreft differentieële effecten voor overachieveers en onderachieveers in de Gesamtschule versus het categoriale systeem. Als een overachiever met een matige intelligentie de sprong naar een hoger categoriaal schooltype kan maken heeft hij grote kans te worden meegenomen door de classesamenstelling en daarvan te profiteren. Een qua intelligentie vergelijkbare onderachiever komt in



het categoriale systeem in de negatieve spiraal van het laagste schooltype en zou beter af zijn geweest in de Gesamtschule. Het lijkt aannemelijk te veronderstellen dat deze processen op het niveau van de school ook op het level van de klas optreden.

De uitkomst van onze analyse betreffende de twee andere hypothesen was in eerste instantie onverwacht en enigszins teleurstellend. We sluiten niet uit dat deze beide onverwachte uitkomsten herleidbaar zijn tot problemen betreffende collineariteit omdat implementatie en taakgerichtheid samenhangen met het gemiddelde niveau van de klas. Bovendien is een effect van "restriction of range" mogelijk aanwezig vanwege de indeling in niveaugroepen en klassen. Van belang is ook de overweging dat het hier gaat om een betrekkelijk gering aantal klassen. Dit kleine aantal impliceert een lage power voor het ontdekken van een effect van klassevariabelen op leerlingvariabelen of voor crosslevel interacties. Het niet vinden van een significant resultaat hoeft dus niet te betekenen dat het gezochte effect ook niet bestaat.

Mogelijk is ook de operationalisering van de implementatie, gemeten via de leerlingpercepties, niet adequaat. Voorts kan een rol hebben gespeeld hebben dat in deze analyse alleen projectklassen zijn betrokken. In al deze experimentele klassen werd het 'model Freudenthal' uitgevoerd. Mogelijk was de variantie in implementatie te klein om effecten te kunnen vinden.

In tweede instantie vonden we toch bij de laagste en de middelste niveaugroepen effecten van de implementatie van het Model Freudenthal op de prestaties van de leerlingen op de natoets (Tabel 4). Bij de hoogste niveaugroep is dit effect afwezig.

Uit onze toepassing blijkt het nut van multilevelanalyse bij curriculumevaluatie. In tegenstelling tot de uitkomst van een deel van het oorspronkelijke onderzoek waarbij, in een regressie-analyse op individueel niveau, een gering effect van de implementatie van het curriculum (volgens de leerlingpercepties) werd gevonden, konden we aanvankelijk niet tot een effect besluiten. In een nadere analyse vonden we evenwel een positief effect van de implementatie, hetgeen overeenkomt met de eerdere bevindingen in een voorgaande studie. Voor de overige conclusies van het oorspronkelijke onderzoek betreffende de positieve effecten van het model Freudenthal in vergelijking met een meer traditionele aanpak verwijzen we naar het genoemde eindrapport.

## LITERATUUR

- Aitkin, M. & Longford, N. (1986). Statistical modelling issues in class effectiveness studies. *Journal of the Royal Statistical Society*, 149, 1-43.
- Beckerman, T.M. en Good, Th.L. (1981). The classroom ratio of high- and low-aptitude students and its effects on achievement. *American Educational Research Journal*, 18, 317-37.
- Boyd, L.H. & Iversen, G.K. (1979). *Contextual analysis: concepts and statistical techniques*. California: Wadsworth Inc.
- Brophy, J. (1986). *Teacher Effects Research and Teacher Quality*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Dar, Y. & Resh N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal*, 3, 357-374.
- Dreeben, R. & Barr, R. (1987). An Organizational Analysis of Curriculum and Instruction. In: M. T. Hallinan, *The Social Organization of Schools*. New York: Plenum.
- Eeden, P. van den, & Saris, W.E. (1984). Empirisch onderzoek naar multilevel uitspraken. *Mens en Maatschappij*, 59, 165-178.
- Fend, H. (1982). *Gesamtschule im Vergleich*. Bilanz der Ergebnisse des Gesamtschulversuch. Weinheim und Basel.
- Fraser, B. J. & Tobin, K. (1989). *Combining qualitative and quantitative methods in the study of classroom learning environments*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Freudenthal, H. (1973a). De niveaus in het leerproces en de heterogene leergroep met het oog op de middenschool. In: *Gesamtschule conferentie 1973*. Amsterdam/ Purmerend: APS/ Muuses.
- Freudenthal, H. (1973b). *Mathematics as an educational task*. Dordrecht.
- Freudenthal, H. (1980). *Weeding and Sowing*. Dordrecht/ Boston, Reidel.



- Goldstein, H. (1987) *Multilevel models in educational and social research*. London: Griffin en New York: Oxford University Press.
- Good, T.L. and Marshall, S. (1984). Do students learn more in heterogeneous groups? In: P.L. Peterson, L.C. Wilkinson and M. Hallinan (eds.): *The Social Context of Instruction*. Orlando: Academic Press.
- Leeuw, J. de & Kreft, G.G. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics*, 11, 57-85.
- Longford, N. (1986). *VARCL-interactive software for variance component analysis. Applications for survey data*.
- Oakes, J. (1985). *Keeping Track, How Schools Structure Inequality*. New Haven and London: Yale University Press.
- Raudenbush, S. & Bryk A.S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59, 1-17.
- Terwel, J. (1984). *Onderwijs maken. Naar ander onderwijs voor 12-16 jarigen*. Harlingen: Flevodruk. Stichting voor Onderzoek van het Onderwijs SVO-reeks nr.77 (proefschrift).
- Terwel, J., Herfs, P., Dekker, R., en Akkermans W. (1988). *Implementatie en effecten van interne differentiatie*. Den Haag: Stichting voor Onderzoek van het Onderwijs (SVO). Selecta reeks.
- Terwel, J. (1989). Real Maths in Cooperative Groups in Secondary Education. In: N. Davidson. *Cooperative Learning in Mathematics* (p. 228-264). Menlo Park: Addison-Wesley.

*Manuscript ontvangen 23-10-1989*

*Definitieve versie ontvangen 13-7-1990*